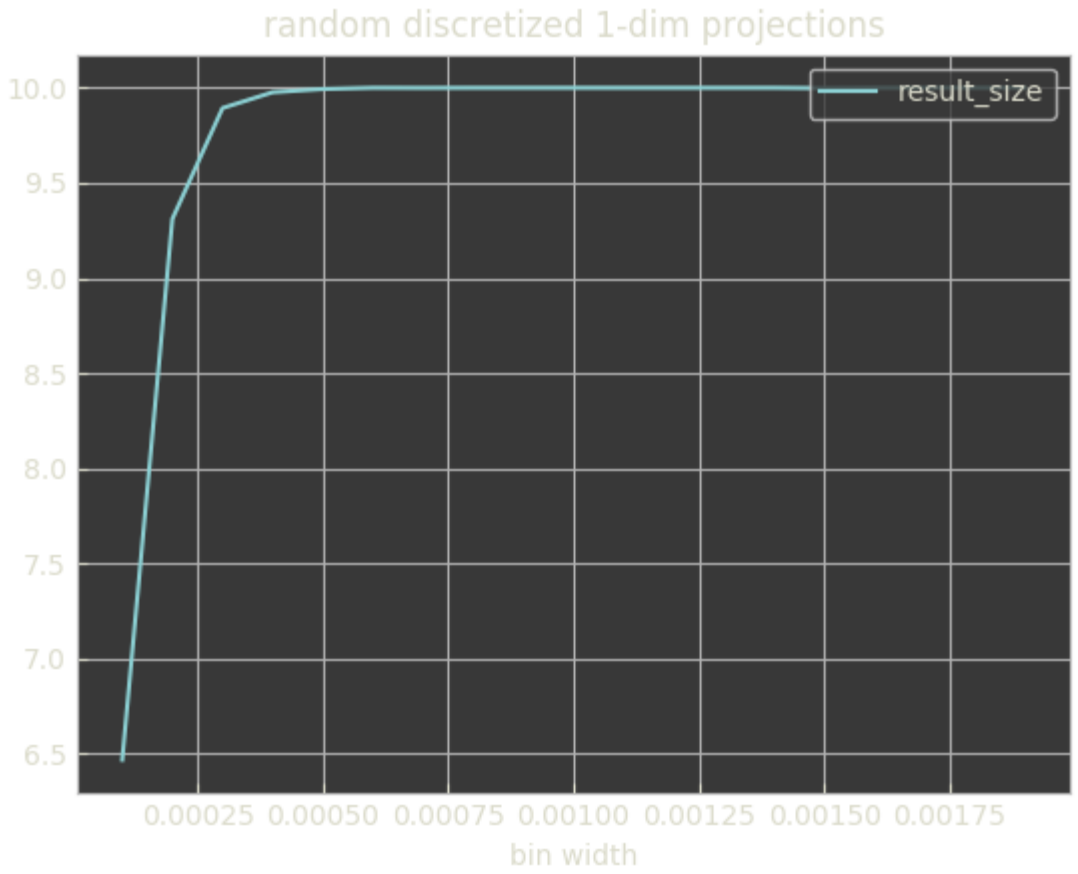
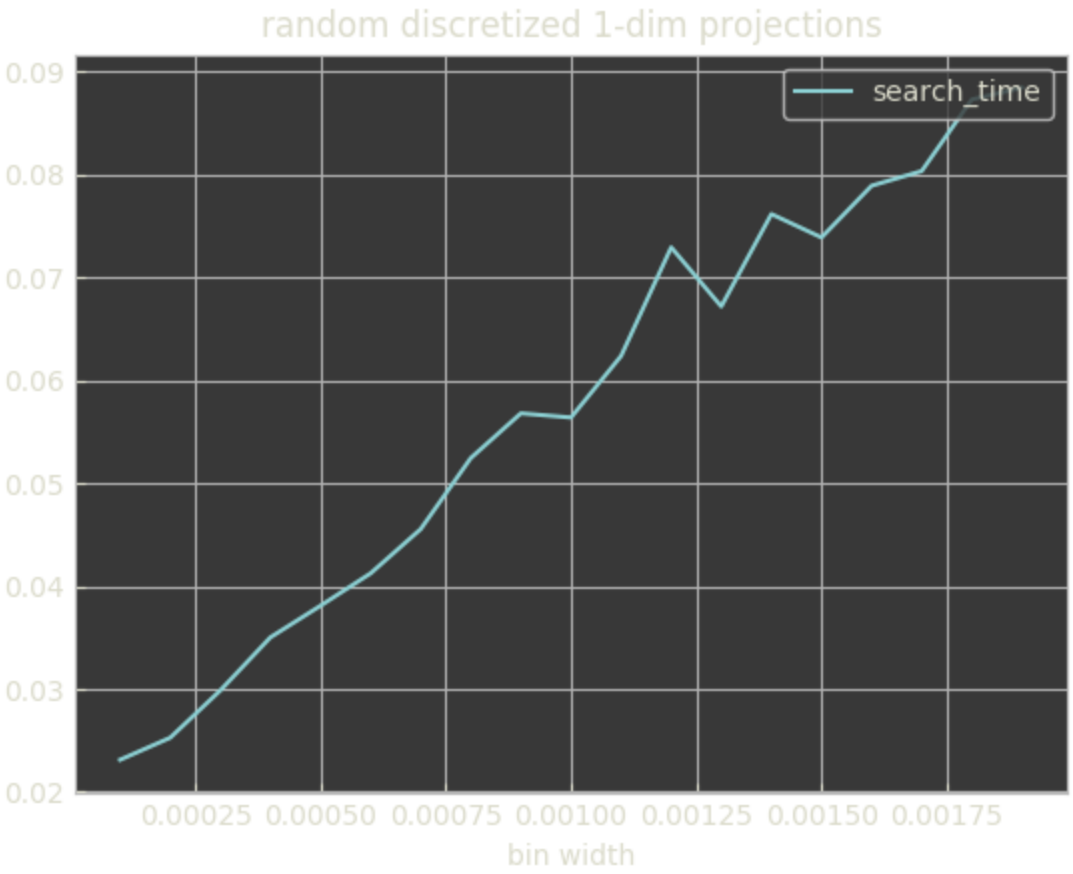
HW 2

1 ANN Algorithm

- NearPy
- Data: Random data, dim = 100, count = 10000, N = 10


random discretized 1-dim projections
result_size
10.0
9.5
9.0
8.5
8.0
7.5
7.0
6.5
0.00025
0.00050
0.00075
0.00100
0.00125
0.00150
0.00175
bin width

- Results:



2 PQ Algorithm

- PQ Fast Scan and PQ Scan
- Data: $\text{ANN}_{\text{SIFT1B}}$
- Results (PQ Fast Scan vs PQ Scan):

  - Mean Instructions Per Query: 554,092,830 vs 118,379,205
  - Mean Run Time: 56,089.95 vs 10,386.41

3 LSH Method

- FALCONN
- Data: Random data, Standard Normal Data with n = 1M, D = 128, 1K Uniform Query on [0, 1M - 1]
- Results:
  - Succes Rate: 0.894 and 0.912 for HP and CP respectively
  - Mean Run Time: 8.26e-4 and 2.54e-4 vs 1.96e-2 for Linear Scan
  - Speed-ups: 23.71% and 77.25%

4 Difficulties

- Most implementations provided by Professor has poor compatibility with gcc 8 and Python 3.7+
- I tackled this with Travis CI and Pyvenv

Author: Lucius Hu

Created: 2019-12-05 Thu 12:03

Validate